EEG Feature Extraction using Parametric and Non-Parametric Models

F. Shiman, S. H. Safavi, F. M. Vaneghi, M. Oladazimi, M. J. Safari, and F. Ibrahim, *Member, IEEE*

Abstract— **We have conducted extensive review on parametric and nonparametric methods for EEG feature extraction and application. We believe that this is the first attempt to compare all methods. Our findings indicate that parametric method does not provide good performance for EEG signal while non-parametric method lack of detail information on the EEG analysis.**

I. Introduction

Exploring brain electrical activities using electroencephalogram (EEG) signals has increased recently. Brain-Computer Interface (BCI), classification of sleep stages, person authentication etc. are some applications of EEG signal analyzing. There are different methods by which features of EEG signals could be extracted and analyzed. These methods generally could be categorized in two basic categories, one of which is called "Non-Parametric Methods" and the other one is the "Parametric Methods".

Non-parametric methods are the most common method used for analyzing EEG signals. In this method, the Gaussian random procedure is detected by statistical possessions of EEG signals in which the signals can be explained through the first and second order moment. Amplitude distribution, interval distribution, correlation analysis etc. are some example using this method.

In Parametric methods, a parametric model is applied to describe the signal which can enhance the estimators[1]. Autoregressive (AR), Moving average (MA) and Autoregressive moving average (ARMA), are some forms by which feature extraction of EEG signals is performed.

In this paper we evaluated 5 methods that were extracted from references [2]-[8].

- Independent Component Analysis (ICA)
- Correlation Analysis
- Power Spectral Entropy
- Autoregressive Modeling
- ARMA Modeling

Manuscript received November4, 2011.This research is supported and funded by the Prime Minister's Department, Malaysia, through a special grant Project No. 66-02-03-0061/H-00000-37039.

F. Shiman, S. H. Safavi, F. M. Vaneghi, M. Oladazimi, M. J. Safari, and F. Ibrahim are with the Medical Informatics and Biological Electro-Mechanical-Systems(MIMEMS),Specialized Laboratory, Department of Biomedical Engineering, Faculty of Engineering, University of Malaya,50603 Kuala Lumpur, Malaysia(e-mail: fatimah@um.edu.my).

II. Methods

Research based on the following references [6]-[14] were selected the advantages and disadvantages of the stated methods are investigated. Finally, by categorizing the non-parametric and parametric methods, the generalized advantages and disadvantages of these two models are obtained. The 5 different EEG feature extraction are expressed as followings:

A. Independent Component Analysis (ICA)

ICA is a method capable of producing subcomponents of a multivariate signal presuming the sources of the signal to be mutually independent and also non-Gaussian. A mathematical formulation representing ICA is given as:

$$x_1(t) = a_{11}s_1(t) + a_{12}s_2(t) + \cdots + a_{1n}s_{1n}(t) \quad (1)$$
$$x_2(t) = a_{21}s_1(t) + a_{22}s_2(t) + \cdots + a_{2n}s_{2n}(t) \quad (2)$$
$$x_k(t) = a_{k1}s_1(t) + a_{k2}s_2(t) + \cdots + a_{kn}s_{kn}(t) \quad (3)$$

Where $x_i(t)$ indicates the k observations and n source signals are illustrated by $s_i(t)$.The combination matrix consists of weight coefficients $\{a_{ij}\}$ with size kn which are related to several undefined parameters (e.g. conduction of modeling volume providing the source localization using scalp electrodes).

ICA is used to estimate the source signals $s_i(t)$ based on the recorded signals $x_i(t)$, with assumptions that the source signals are non-Gaussian and statistically independent[2]. However, in some cases which the location of signal sources is needed, ICA is not effective. Nevertheless, some solutions have been proposed to tackle this problem [2].

ICA has applications in many areas and was reported to give excellent performance in EEG analysis. An EEG data are recorded by electrodes which are electrical potentials in many special locations on the scalp and it is generally accepted that unidentified components of neural source activity are extracted from these scalp recordings with linear combinations [2].

B. Correlation Function

Investigating the correlation among different randomly selected variables can be achieved through a correlation function. It is demonstrated as a function of temporal or spatial distance between two specific points.

Correlation functions of various random variables are sometimes identified as cross correlation functions. This is to highlight that different variable are significant as they are created of cross correlations.

978-1-4577-2177-9/12/$25 (C) 2012 IEEE

A correlation function can certainly referred to as one of the implications for a novel spectral analysis of EEG. Random data for correlation function expresses the significances of the data at the same time on the significance of the same data in autocorrelation analysis in terms of the common dependence. This equation can describe the cross correlation for x and y signals as

$$\Phi_{xy}(\tau) = E\{x(t)y(t + \tau)\} \qquad (4)$$

Where τdenotes the lag time. Providing the condition of x=y, $\Phi_{xy}(\tau)$ represents the function of autocorrelation and also can be calculated for information of discrete values [1].

This method may be used for estimation of parametric model [3]. The estimation in this method is unbiased and inconsistent[4]. Polarity coincidence correlation function [5], auto- or cross-averaging[6, 7] and complex demodulation [8] are some modifications which have been used for this method.

Correlation analysis could be applied for monitoring ischemic changes of EEG and researches show that the high precision can discover the ischemic event although it has a delay in data computation [9]-[10]. Despite many advances achieved in this method application, many neuroscientists still confirmed the cross-correlation between the performance of pairs of neural constructions to assume their functionality [11].

C. Power spectral entropy (PSE)

PSE measures the spectral complication of an uncertain system through information entropy. It assumes a random variable X as shape of the system for an uncertain system, and X were describe as

$$\mathrm{X} = \{\mathrm{x}_1, \mathrm{x}_2, \dots, \mathrm{x}_\mathrm{n}\} \qquad (\mathrm{n} \geq 1)$$

The corresponding probability is

$$\mathrm{P} = \{\mathrm{p}_1, \mathrm{p}_2, \dots, \mathrm{p}_\mathrm{n}\} 0 \leq \mathrm{p}_\mathrm{i} \leq 1, \mathrm{i} = 1,2, \dots, \mathrm{n}$$

Under Constraints

$$\sum_{\mathrm{i}=1}^{\mathrm{n}} \mathrm{p}_\mathrm{i} = 1 \qquad (5)$$

Consequently, the information regarding entropy of the method can be represented as follows:

$$\mathrm{H} = -\sum_{\mathrm{i}=1}^{\mathrm{n}} \mathrm{p}_\mathrm{i} \ln \mathrm{p}_\mathrm{i} \qquad (6)$$

Fast Fourier Transform (FFT)coverts the time-series of signals into the power spectrum referred to as information entropy of power spectrum as power spectral entropy [12].

PSE is described by a quantity of time indecision in frequency domain. The entropy value of power spectrum in EEG signals is small when the spectrum peak is narrow. It illustrates an apparent concussive rhythm in the signal that results to a low complication when the wave is in order. While having smoother peak of the spectrum, the entropy has a greater value and as a consequence, spectra structure of EEG signals can be revealed by PSE [12]. Many alterations were applied in these methods to obtain consistent estimates [1].

EEG signals have been categorized into various frequency bands in terms of frequency contents. The brain dysfunction implies for having components of power and frequency within these specific bands. Regarding the spectral characteristics of EEG signal, time dependent variables of power spectra have been taken into consideration for further analysis of time variations[13].One of the most significant benefits of this type of analysis is related to maintain all information contents of EEG when it transforms the artifacts with low frequencies into a narrow frequency spectrum [1].

In contrast, the disadvantages of these methods related to limitation of feature analysis in terms of labor intensiveness, inter-user variability, and storage problems. They also create linear and grayscale displays of spectral analysis to decrease the labor intensiveness [14],which can further define the sum of the presented data.

D. Autoregressive (AR)

AR modeling is one of the prominent parametric methods. It indicates that linear mixture of the past EEG samples plus an independent component (white noise) brings existing EEG sample.

The forward prediction of the EEG signal was accomplished using the following equation:

$$x[n] = \sum_{i=1}^{p_t} a_i x[n-i] + e[n] \qquad (7)$$

Where e[n] described as prediction error (new information contained in the current EEG sample) [15].

Several features expressed the reason for reputation of AR modeling of EEG: i) Short-term EEG spectrum can be distinguished by AR process with sensible accuracy; ii) AR model is totally applied in time series analysis context; iii) Parameters of AR model are estimated by simple algorithms. As expected, AR models are suitable choice to analyze EEG for biomedical engineers [15].

Linearly dependency of this model to past values, which is common in modeling approaches, may decrease accuracy of this method. However, Xuan Kong [15] increased the accuracy by adding a predictive part to above mentioned equation.

E.ARMA Modeling

An ARMA model is also one of the parametric models including an autoregressive (AR) part and a moving-average (MA) part. In addition it is utilized as a predictor for time series feature values and also predicts at special time instance according to random distribution and past value(AR and MA parts) [16].

ARMA (m,n) model is:

$$\mathrm{y}_\mathrm{t} = -\sum_{\mathrm{j}=1}^{\mathrm{m}} \mathrm{a}_\mathrm{t}^{(\mathrm{j})} \mathrm{y}_{\mathrm{t}-\mathrm{j}} + \sum_{\mathrm{k}=1}^{\mathrm{n}} \mathrm{b}_\mathrm{t}^{(\mathrm{k})} \mathrm{e}_{\mathrm{t}-\mathrm{k}} + \mathrm{e}_\mathrm{t} \qquad (8)$$

y_t represents time sample of the EEG signal corresponding to a single channel, $a_t^{(j)}$ is autoregressive and $b_t^{(k)}$ is moving average parameter at discrete time instant *t* and n, m are the zeros and poles numbers respectively. White noise Gaussian is defined by e_t [16].

When frequency spectrum demonstrates both sharp peaks and deep nulls, this method is the most excellent model processes [16]. We can generate time-frequency spectra in parametric model with time varying parameters as the resolution is higher than short-time Fourier transform or Wavelet generated spectra. In majority of EEG applications ARMA generally referred to as a more common representation of the AR; regarding the fact that EEG signals seems to correspond more to this model, it would be more sufficient to apply such an accurate model [17].

III. Performance Of Methods

For evaluating the performance of different EEG feature extraction methods, first of all, we should define the term "performance" and determine its indications. If we define method "performance" with indications as higher accuracy, precision and speed, then it is required to determine the application of method as well. In other word, higher speed could be considered as an indication of performance if the method is used in an EEG monitoring system, whereas higher accuracy is more crucial for BCI applications rather than EEG monitoring purposes. It is to say that, performance of different EEG feature extraction method could be measurable if specific application is desired.

As it was mentioned earlier, performing an extensive investigation regarding the 5 different methods with the same application did not lead to any specific study. Some studies were found in which performance evaluation of specific EEG analyzing algorithm were conducted in comparison with fMRI modality as a control/gold standard [18]. These studies were not enough to make a comparison table between five different methods.

In general terms, we can divide selected feature extraction methods into two categories of non-parametric and parametric. In consequence, correlation analysis and power spectral entropy will be placed in first category and ICA, Autoregressive Modeling and ARMA Modeling will be considered in the second category.

As it was mentioned before, ICA assumes EEG signals as a linear summation of several independent signal sources. This property makes ICA a proper choice for person authentication and identification applications to extract signal sources despite the EEG electrodes position. However, the number of electrodes must be assumed to be equal or more than the sources [2]. This means ICA is unable to identify the actual number of source signals. Thus, it is highlighted in blind signals.

The correlation function for random data describes the general dependence of the values of the data at one time on the values of the same data in the case of autocorrelation analysis (or of different data in the case of cross-correlation analysis) at another time [1]. However, in 1993, Westdorp conducted an investigation [19] showing that the EEG signal is normally correlated when compared with Laplacian approach. The popularity of this method has reduced although there are solutions to deal with this problem.

As it was mentioned earlier, AR is the most popular parametric method to analyze EEG signals. It provides more details on spectrum data in comparison with non-parametric methods. Salleh *et al*. [20] investigated the AR model for the EEG signal analysis during Salat meditation. In their study the spectral AR was compared with FFT and the finding revealed that the spectral AR performed better than the FFT. FFT gave poor result due to its spectral leakage. The advantages of using AR technique include smoother and more easily interpretable power spectrum. However, Hosni, S.M. *et al*. [21] has proven that the best EEG classification accuracy is AR model in comparison with AR spectral analysis and power differences. This AR model was found to be the most suitable for clinical applications [17]. Thus, in our study, we will use AR model to analyze the EEG signals namely alpha and Gamma bands during Salat meditation.

The main drawback of spectral analysis is that it needs a longer observation time to achieve optimum spectral estimation. This will then cause a conflict to the non-stationary behavior of the EEG signals. Another drawback is that it's hard to get the desired end result, since certain important values such as bandwidths, peak frequencies and fractional power quantities were not provided. If the values were calculated from the power spectrum, then there is no guarantee that the estimators will be efficient. These drawbacks are essentially eliminated by applying parametric models. Parametric models indicate a considerable changes of the spectral properties in examples of placebo influenced EEG's which were not detected by visual assessment of the EEG [22].

However, the defined models are descriptive and empirical, and consequently cannot reconstruct whole neurophysiologic specifications of the EEG. The defined model produces particular linear signal character because the higher order coefficients were not considered. Thus, it's not appropriate to choose this type of model when it comes to amplitude distributions evaluation. Nevertheless, the model is quite adequate as long as spectral analysis is the main concern.

In general, linear analysis schemes, which were discussed in this paper, only utilize information retained in the autocorrelation function (i.e., the second-order cumulant). Additional information stored in higher-order cumulants is therefore ignored by assumption. Thus, while the power spectrum provides the energy distribution of a stationary process in the frequency domain, it cannot distinguish nonlinearly coupled frequency from spontaneously generated signals with the same resonance condition [23].Table1.shows the summary of the advantages and disadvantages of above mentioned methods and their main applications to make it easier to compare the performances.

TABLE I
SUMMARY OF METHODS

Method Name	T	Advantages	Disadvantages	Main Application
ICA	Parametric	• Proper choice for person authentication and identification applications.	• Unable to identify the actual number of source signals as well as their locations • It is widely based on linearity	• Person authentication and identification
Correlation Analysis	Nonparametric	• This method may be used for estimation of parametric model. • The estimation in this method is unbiased.	• EEG signal is fairly correlated normally in comparison with Laplacian analysis • The estimation in this method is not consistent • Delay for computation of Data.	• It is used for monitoring ischemic changes of EEG by high precision. • Delay for computation of data.
PSE	Nonparametric	• It has good usefulness in classification of functions and dysfunctions of EEG signals based on frequencies and power • Reflects the spectra structure of EEG signals	• It does not have good usefulness whenever the predictable pattern of EEG is desirable • Not good for monitoring • No guarantee that the estimates will be efficient for certain characteristic values like peak frequencies, bandwidths, and fractional power quantities.	• Good for classification of signals • This method seems to be more common to extract EEG feature.
AR	Parametric	• Linearly dependency of this model to past values, which is common in modeling approaches, may decrease accuracy of this method • The best classification accuracy between AR, AR spectral analysis and power differences is AR model. • The model based scheme is far superior to some non-parametric approaches as quadratic detection filter.	• Distinguished the short-term EEG spectrum • Totally studied in time series analysis context • It can provide more detail on spectrum data in comparison with non-parametric methods. • The model is insufficient to explain properties of higher order correlation coefficients for EEG signal • Inappropriate to use when amplitude distributions are discussed	• Useful in real-time estimation • Suitable choice to analyze EEG for biomedical engineering's • For Spectral analysis, the model is quite adequate.
ARMA Modeling	Parametric	• The most efficient method for modeling courses having both sharp peaks and deep nulls within their frequency spectrum. • The model based scheme is far superior to some non-parametric approaches as quadratic detection filter	• It is inadequate for Characterizing the coefficients of high orders being applied to EEG signal analysis. • Even though ARMA is the more generalized representation of the AR modeling, still AR is well suited for most of EEG applications.	• Useful in real-time estimation

IV. CONCLUSION

Two well-known non-parameterized and parameterized models identified in this paper are the spectrum analysis and Autoregressive (AR) methods, respectively[22]-[25]. In this paper, it is very difficult to discuss and to make comparison on all methods capabilities as shown in Table 1. The findings summarized that each method has specific advantages and disadvantages which make it suitable for a specific application. Parametric methods, which assume a predefined pattern for EEG signals, may not provide high-quality performance for some EEG signals. In contrast, non-parametric methods, for instance, may not provide detail information on EEG analysis as much as parametric methods. Therefore, the performance of the methods used will depend on the specific EEG application.

REFERENCES

[1] C. De Luca, "Encyclopedia of Medical Devices and Instrumentation," *John G. Webster,* 2006.

[2] C. He and J. Wang, "An Independent Component Analysis (ICA) Based Approach for EEG Person Authentication," pp. 1-4.

[3] P. S. S. D. Gannabathula and I. Murthy, "ARMA model estimation for EEG using canonical correlation analysis," pp. 1204-1205 vol. 3.

[4] A. B. Donahue, "Electroconvulsive therapy and memory loss: a personal journey," *The journal of ECT,* vol. 16, p. 133, 2000.

[5] R. Abrams, "Does brief-pulse ECT cause persistent or permanent memory impairment?," *The journal of ECT,* vol. 18, p. 71, 2002.

[6] A. Kamp, W. S. van Leeuwen, and A. Tielen, "A method for auto-and cross-relation analysis of the EEG," *Electroencephalography and Clinical Neurophysiology,* vol. 19, pp. 91-95, 1965.

[7] A. Remond, N. Lesèvre, J. Joseph, H. Rieger, and G. Lairy, "The alpha average. I methodology and description," *Electroencephalography and Clinical Neurophysiology,* vol. 26, pp. 245-265, 1969.

[8] R. Dean Linden, T. W. Picton, G. Hamel, and K. B. Campbell, "Human auditory steady-state evoked potentials during selective attention," *Electroencephalography and Clinical Neurophysiology,* vol. 66, pp. 145-159, 1987.

[9] L. Czinege, R. Urbanics, and Z. Farkas, "Multi-channel activity correlation analysis-a method to detect ischemic changes of the EEG," 1995, pp. 735-736 vol. 1.

[10] L. Czinege, R. Urbanics, and Z. Farkas, "Multi-channel EEG activity correlation analysis to detect the onset of cerebral ischemia," 1994, pp. 1230-1231 vol. 2.

[11] L. A. Baccalá and K. Sameshima, "Overcoming the limitations of correlation analysis for many simultaneously processed neural structures," *Progress in brain research,* vol. 130, pp. 33-47, 2001.

[12] A. Zhang, B. Yang, and L. Huang, "Feature extraction of EEG signals using power spectral entropy," 2008, pp. 435-439.

[13] F. Lopes da Silva, "EEG analysis: theory and practice," *Electroencephalography: Basic Principles, Clinical Applications and Related Fields, 4th edition,* pp. 1135-1163, 1998.

[14] W. J. Levy, H. M. Shapiro, G. Maruchak, and E. Meathe, "Automated EEG processing for intraoperative monitoring: a comparison of techniques," *Anesthesiology,* vol. 53, p. 223, 1980.

[15] X. Kong, "Forward and backward autoregressive modeling of EEG," pp. 1215-1217 vol. 3.

[16] T. Cassar, K. Camilleri, S. Fabri, M. Zervakis, and S. Micheloyannis, "ARMA modeling for the diagnosis of controlled epileptic activity in young children," pp. 25-30.

[17] C. Kellner, D. Nixon, and H. Bernstein, "ECT--drug interactions: a review," *Psychopharmacology bulletin,* vol. 27, p. 595, 1991.

[18] M. J. Barton, *et al.*, "Evaluating the performance of kalman-filter-based eeg source localization," *Biomedical Engineering, IEEE Transactions on,* vol. 56, pp. 122-136, 2009.

[19] A. Westdrop, "Volume conduction effects on correlation analysis of EEG data," 1993, pp. 150-152.

[20] N. Salleh, K. Lim, and F. Ibrahim, "AR modeling as EEG spectral analysis on prostration," 2009, pp. 1-4.

[21] S. M. Hosni, M. E. Gadallah, S. F. Bahgat, and M. S. AbdelWahab, "Classification of EEG signals using different feature extraction techniques for mental-task BCI," 2007, pp. 220-226.

[22] A. Isaksson, A. Wennberg, and L. Zetterberg, "Computer analysis of EEG signals with parametric models," *Proceedings of the IEEE,* vol. 69, pp. 451-461, 1981.

[23] J. D. Bronzino, *The biomedical engineering handbook* vol. 2: CRC Pr I Llc, 2000.

[24] A. Rémond, I. F. o. S. f. Electroencephalography, and C. Neurophysiology, *EEG Informatics: A Didactic Review of Methods and Applications of EEG Data Processing: Lectures for an*

International Course in EEG Data Processing, Held Under the Auspices of the International Federation of Societies for Electroencephalography and Clinical Neurophysiology in Paris on August 29-September 2, 1977: Elsevier Scientific, 1977.
[25] B. H. Jansen, J. R. Bourne, and J. W. Ward, "Autoregressive estimation of short segment spectra for computerized EEG analysis," *Biomedical Engineering, IEEE Transactions on,* pp. 630-638, 1981.